

Sains Malaysiana 39(4)(2010): 677–684

Skor Pengukuran Kebergunaan Perisian Kursus Matematik Berdasarkan Faktor Penilai

(Mathematics Courseware Usefulness Measurement Score
Based on Evaluator's Factors)

NORAI DAH SAHARI @ ASHAARI*, HAIRULIZA MOHD. JUDI, ABDUL AZIM ABDUL GHANI,
MOHD HASAN SELAMAT & AIDA SURAYA MD. YUNUS

ABSTRAK

Tujuan utama kajian ini adalah untuk menghasilkan bukti kesahan dan kebolehpercayaan metrik pengukuran kebergunaan perisian kursus matematik secara psikometrik dan seterusnya membentuk rumus skor kebergunaan. Kertas ini membincangkan pembentukan rumus kebergunaan. Berdasarkan pandangan secara teori, satu model hipotesis kebergunaan dengan tiga faktor iaitu kebolehgunaan, kefungsihan dan kecekapan dan 85 metrik dicadangkan. Melalui pengesahan pakar dan ujian kebolehpercayaan metrik kebergunaan dikurangkan kepada 64. Instrumen dengan 64 metrik ditadbir ke atas 620 guru matematik di lima zon di Malaysia. Mereka diminta melayari perisian kursus matematik (PKM) sambil melengkapkan instrumen. Analisis faktor penjelajahan dilakukan untuk mengesahkan konstruk kebergunaan manakala analisis faktor pengesahan dijalankan untuk membentuk model berstruktur dengan memadankan 34 metrik dengan model hipotesis. Ujian-t dan ujian ANOVA mendapati tiga pembolehubah iaitu, bidang pengkhususan guru, pengalaman menggunakan komputer dan ralat penilaian menyumbang kepada skor kebergunaan. Dapatan ini digunakan untuk membentuk rumus kebergunaan PKM.

Kata kunci: Metrik kebergunaan; model kebergunaan; pengesan ralat; ralat penilaian; skor kebergunaan

ABSTRACT

The primary purpose of this study was to provide psychometrics evidence of validity and reliability of metrik for evaluating the usefulness of a mathematics courseware (MC) and eventually to construct the usefulness score formula. In this paper the construction of usefulness score is discussed. Based on the theoretical perspective, a hypothesized evaluation model with three factors which were usability, functionality and efficiency and 85 metrik were proposed. The instrument was administered to 620 mathematics teachers in five regions in Malaysia. They were required to explore a MC and complete the questionnaire. Exploratory factor analysis was employed to validate the usefulness construct while confirmatory factor analysis was conducted to develop a structural model by befitting 34 metric with the hypothesis model. T-test and ANOVA suggested that three variables, which were teacher's academic background, computer usage experience and evaluation error, had significant impact to MC usefulness score. From these findings MC usefulness score formula was constructed.

Keywords: Evaluation error; safeguards; usefulness metrik; usefulness model; usefulness score

PENDAHULUAN

Subjek matematik merupakan mata pelajaran yang wajib diambil oleh semua pelajar sekolah. Memandangkan matematik merupakan suatu subjek yang sukar dikuasai dan memerlukan kaedah pembelajaran yang konkrit, penggunaan pakej pembelajaran berbantuan komputer dapat mengatasi masalah pembelajaran subjek ini. Menurut Papert (1998) penggunaan perisian kursus (PK) dapat menyediakan pengalaman pembelajaran yang konkrit dan pelbagai aktiviti yang melibatkan pembinaan pengetahuan seperti penyelesaian masalah, membuat eksperimen dan penjelajahan berkaitan dengan topik dapat dilakukan dengan bantuan komputer. Dengan kemajuan teknologi multimedia kebanyakan aktiviti

pembelajaran matematik dapat diintegrasikan di dalam satu pakej PK yang dapat menyamai aktiviti pembelajaran di bilik darjah (Ministry of Education Malaysia 1997). Justeru, banyak pakej perisian kursus matematik (PKM) telah dibangunkan.

PKM memberi kemudahan persekitaran bagi penjelajahan dan penemuan dan pelajar dapat membina pengetahuan mereka melalui aktiviti berinteraksi dengan komputer. Maklum balas atau respon serta merta yang disediakan meningkatkan keyakinan dan motivasi. Ia juga meningkatkan kemahiran menyelesaikan masalah memandangkan PKM ini mampu menyediakan pelbagai masalah berkaitan kehidupan harian. Malah kekuatan yang dihasilkan oleh teknologi komputer ini adalah

keupayaannya membawa realiti ke dalam persekitaran pembelajaran dan disokong dengan alat pengukuran dan alat analisis (Yerushamly & Schwartz 1999).

Menurut Kimmins (1995), teknologi berupaya memberi kuasa kepada pelajar untuk bermain dengan pelbagai perwakilan, meningkatkan keupayaan membuat visualisasi dan bertanggung jawab terhadap pembelajaran mereka. Quintana et al. (2001) pula menambah, teknologi multimedia dapat membantu dalam pendidikan matematik dengan memperoleh, menilai dan memproses maklumat numerik, menjalankan pengiraan, menggunakan grafik untuk berkomunikasi dengan maklumat numerik, menyiasat dan menyelesaikan masalah, mereka bentuk model dan simulasi serta memberi bimbingan bagi konsep aras tinggi yang abstrak. Yushau et al. (2004) mendapati guru dan pelajar telah menunjukkan sikap positif terhadap beberapa PKM. Pelajar percaya mereka mendapat faedah daripada pembelajaran menggunakan teknologi komputer (Atkins 2005). Keupayaan teknologi ini dapat dimanfaatkan dengan sebaiknya jika pembangunan PKM dilakukan dengan baik.

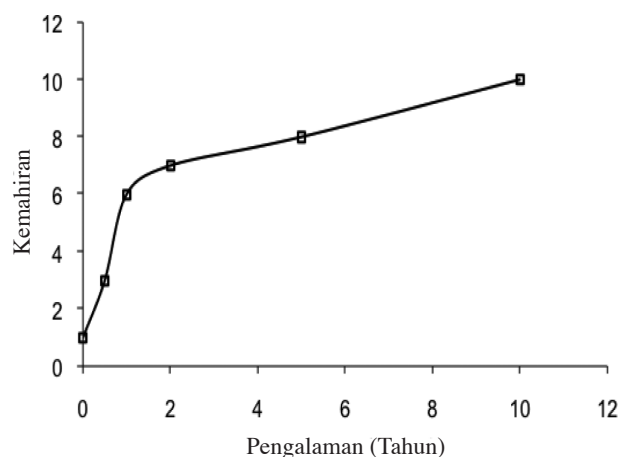
Memandangkan penggunaan PKM dalam pengajaran dan pembelajaran (P&P) didapati berkesan, banyak PKM telah dibangunkan. Sering kali pembangunan sesebuah perisian hanya memfokus kepada nilai komersial iaitu kualiti media, bukan elemen yang meningkatkan kualiti P&P serta perubahan tingkah laku pelajar (Jamalludin & Zaidatun 2003). Bila banyak PKM berada di pasaran, pemilihan perlu dilakukan agar PKM yang terbaik dan memenuhi keperluan kurikulum dan mudah digunakan (ease of use) diperoleh. Bagi mengatasi masalah ini campur tangan orang yang berkepentingan seperti guru amat perlu. Mereka seharusnya dapat membuat keputusan sama ada perisian yang dibangunkan itu menepati kehendak pengguna khususnya pelajar dan sama ada PK yang dibangunkan itu menepati objektif pembelajaran matematik.

Satu instrumen diperlukan untuk memastikan yang suatu PKM itu menepati keperluan pendidikan dan penggunaannya berkesan. Instrumen tersebut seharusnya dapat digunakan oleh mereka yang mahir dan perihatin dengan keperluan pelajar iaitu, guru. Untuk itu, bagi membantu guru membuat pilihan, penyelidik membangunkan satu instrumen bagi menilai atribut kualiti kebergunaan PKM sebelum digunakan. Kebergunaan merupakan satu atribut kualiti yang mengukur sama ada suatu PKM itu dapat memberi manfaat kepada pelajar, dan mengikat minat pelajar (engage) menggunakannya sebagai bahan pembelajaran sendiri. Bagi memastikan pengukuran kebergunaan yang dijalankan adalah betul, kaedah pengiraan skor penilaian yang tepat diperlukan.

SKOR PENILAIAN

Pembangunan instrumen bertujuan untuk mengukur dan menambah baik kualiti produk yang dinilai. Lazimnya, pengiraan skor bagi penilaian menggunakan soal selidik adalah dengan mengira purata maklum balas penilaian.

Xenos dan Christodoulakis (1995a) telah menggariskan panduan untuk menentukan penilaian berasaskan pandangan pengguna dengan mengambil kira kelayakan dan pengalaman penilai serta ralat penilaian dalam pengiraan min skor soal selidik. Atribut pengguna yang tidak berkaitan dengan produk yang dinilai dikumpulkan di bawah kategori latar belakang peribadi, misalnya umur, jantina dan tahap pendidikan. Pemboleh ubah ini dapat menentukan kelayakan penilai. Kekerapan dan pengalaman menggunakan komputer dalam melakukan tugas harian serta menggunakannya dalam pengajaran dimasukkan ke dalam kategori pengetahuan pengguna. Menurut Xenos dan Christodoulakis (1995a), pengguna mendapat pengalaman lebih cepat pada tahun pertama menggunakan suatu program dan bertambah secara perlahan selepas itu seperti yang digambarkan oleh graf Rajah 1.



RAJAH 1. Graf pengalaman dan tahun menggunakan produk perisian

Pengukuran kemahiran atau pengalaman pengguna mengoperasikan suatu produk perisian dapat menilai kebolehan mereka dalam memberi kritikan yang objektif terhadap perisian yang serupa. Berpanduan kepada Rajah 1, Xenos dan Christodoulakis (1995a) menerangkan bahawa tiga daripada sepuluh pengalaman diperoleh selepas setengah tahun menggunakan suatu perisian, dan enam daripada sepuluh pengalaman pada keseluruhan tahun pertama. Pengguna secara seragam mencapai sepuluh kemahiran selepas sepuluh tahun menggunakan perisian. Selain daripada kelayakan dan pengalaman penilai, satu lagi ciri penting yang mempengaruhi penilaian berasaskan pendapat pengguna suatu perisian adalah ralat penilaian.

RALAT PENILAIAN

Sudah menjadi kebiasaan di dalam kajian tinjauan, kesilapan dalam memberi maklum balas sering terjadi. Maklum balas yang diberi ada kalanya tidak melambangkan pendapat penilai. Mungkin penilai tidak menilai secara serius atau jujur yang menyebabkan ralat berlaku. Menurut

Xenos & Christodoulakis (1999), ralat ini mungkin berlaku disebabkan beberapa perkara seperti berikut:

1. Penilai cuai semasa menjawab soal selidik dan menanda secara rawak apabila mereka keliru atau tidak membaca arahan.
2. Penilai menunjukkan minat positif untuk menjawab soal selidik pada peringkat permulaan sahaja tetapi hilang minat apabila di pertengahan soal selidik lalu menjawab sambil lewa.
3. Penilai berminat menjawab keseluruhan soal selidik tetapi mungkin keliru menyebabkan tidak sengaja memberi jawapan yang salah.

Ralat tersebut dapat dikurangkan jika soal selidik disediakan dengan arahan yang jelas dan metrik mudah difahami. Namun begitu, ralat penilaian tidak boleh dihapuskan begitu sahaja. Untuk mengendalikan jawapan soal selidik yang mengandungi ralat, penggunaan pengesan ralat (safeguard) adalah wajar. Xenos dan Christodoulakis (1999) telah mencadangkan penggunaan pengesan ralat untuk mengukur sama ada penilai menilai secara jujur atau tidak. Terdapat tiga jenis pengesan ralat:

1. Soalan perangkap: Soalan jenis ini memerlukan satu jawapan muktamad oleh responden tertentu. Maklum balas yang lain menunjukkan ralat.
2. Soalan yang diulang dengan ayat yang berbeza: Jawapan yang berlainan bagi maksud yang sama menunjukkan ralat penilai.
3. Soalan yang sama dengan pilihan jawapan yang berlainan: Memilih jawapan yang tidak tekal menunjukkan ralat penilai.

Dengan menggunakan pengesan ralat, penilai yang menjawab soal selidik secara tidak serius boleh dikenal pasti semasa data dianalisis. Namun begitu, terlalu banyak pengesan ralat akan memberi kesan terhadap konsistensi soal selidik. Bilangan pengesan ralat yang dicadangkan adalah 5% hingga 10% daripada jumlah metrik (Xenos & Christodoulakis, 1995(b)). Bahagian seterusnya menerangkan metodologi kajian untuk menghasilkan rumus kebergunaan.

METODOLOGI KAJIAN

Bahagian ini menerangkan tentang reka bentuk pembangunan dan pengesanan instrumen, peserta kajian dan prosedur pengumpulan data secara ringkas. Di sepanjang penyelidikan ini jumlah peserta kajian adalah 668 orang dan enam jenis PKM digunakan. Setiap peserta perlu melakukan pemeriksaan heuristik terhadap PKM dan melengkapkan instrument pengukuran kebergunaan PKM (IPKpkm). Kaedah pemeriksaan heuristik adalah satu kaedah dalam kejuruteraan kebolehgunaan bagi mengenal pasti masalah kebolehgunaan suatu reka bentuk antara muka pengguna supaya suatu perisian dapat dibaiki dalam proses pembangunan yang berulang. Pemeriksaan ini merupakan suatu kaedah yang diadaptasi daripada

pemeriksaan heuristik kebolehgunaan oleh Nielsen (1994). Ia merupakan satu teknik penilaian yang berkesan daripada segi kos dalam menghapuskan kesilapan yang serius di dalam suatu perisian.

Walaupun kaedah heuristik direka dalam konteks kajian kebolehgunaan bagi melihat masalah kebolehgunaan antara muka pengguna, namun Quinn (1996), Squires dan Preece (1996), Albion dan Gibson (1998) dan Diaz (2003) telah mencadangkan teknik yang serupa bagi menilai PK. Kajian heuristik yang dijalankan bukan sahaja memeriksa kebolehgunaan antara muka pengguna, tetapi penilai juga memeriksa kesesuaian isi kandungan, strategi berarahan serta elemen lain yang berkaitan dengan PK. Secara keseluruhannya kajian ini mengandungi tiga fasa.

Dalam fasa pertama metrik penilaian dikumpulkan berdasarkan analisis pemeriksaan heuristik terhadap satu perisian kursus matematik iaitu PKM A oleh lima orang pembangun PK dan analisis ulasan penilaian PKM (Cook 2002; Hawkes 2000; Lawson 2005) serta instrumen sedia ada yang berkaitan. Metrik tersebut membentuk senarai semak IPKpkm_1. Senarai semak ini disahkan metriknya oleh lapan orang penilai pakar dengan menjalankan pemeriksaan heuristik ke atas tiga perisian kursus yang berlainan iaitu PKM B, C, dan D. Seterusnya soal selidik IPKpkm_2 dengan lima skala Likert dibangunkan. Soal selidik IPKpkm_2 ini disemak formatnya dan menghasilkan IPKpkm_3. Seterusnya analisis metrik dengan ujian kebolehppercayaan dan diskriminasi ke atas data penilaian oleh 35 orang guru matematik dilakukan untuk membentuk IPKpkm_4. Instrumen ini digunakan pada fasa ketiga untuk menguji kesahan konstruk, serta pembentukan rumus kebergunaan. Metodologi ini diringkaskan dalam Jadual 1.

Tinjauan oleh 620 orang guru dianalisis dan skor kebergunaan ditentukan menggunakan perisian SPSS 12.0 dan AMOS 5. Ujian hipotesis bagi menentukan pemboleh ubah demografi yang mempengaruhi pemberian skor kebergunaan dijalankan. Ujian-t dilakukan bagi menentukan sama ada skor kebergunaan dipengaruhi oleh jantina responden, pengalaman membangunkan PKM dan sama ada penilai yang melakukan ralat mempengaruhi pemberian skor. Ujian ANOVA satu hala digunakan untuk menentukan sama ada skor kebergunaan dipengaruhi oleh kelulusan akademik dan bidang pengkhususan. Tinjauan ini juga dijalankan bagi melihat sama ada skor kebergunaan dipengaruhi oleh pengalaman menggunakan aplikasi komputer, pengalaman mengajar matematik dan pengalaman menilai PKM. Pembentukan rumus kebergunaan dilakukan berdasarkan skor penilaian sebagaimana yang dicadangkan oleh Xenos dan Christodoulakis (1995b): Jika tiada faktor demografi mempengaruhi penilaian berdasarkan analisis di atas, rumus yang lazim digunakan adalah:

$$\text{Min skor}_p = \frac{\sum_{i=1}^n P_i}{n}, \quad (1)$$

JADUAL 1. Ringkasan metodologi kajian

Fasa	PKM	Instrumen	Bilangan Peserta	Aktiviti dan Analisis
Fasa I	PKM A	Soal selidik soalan terbuka	5 orang	Pemeriksaan heuristik Pembentukan senarai semak IPKpkm _1
Fasa II	PKM B, C dan D	IPKpkm _1	8 orang pakar	Pemeriksaan heuristik Pengesahan metrik Pembentukan soal selidik IPKpkm _2
		IPKpkm _2	4 orang pakar	Semakan soal selidik hasilkan IPKpkm _3
	PKM E	IPKpkm _3	35 orang	Analisis metrik hasilkan IPKpkm _4
Fasa III	PKM E	IPKpkm _4	620 orang	Pemeriksaan heuristik Kesahan Konstruk: Analisis Faktor Penjelajahan, Analisis Faktor Pengesahan, Pembentukan rumus kebergunaan: Ujian-t dan ANOVA

dengan $P_i \in [0.100]$, adalah peratus skor pendapat penilai i bagi PKM yang dinilai dan n adalah bilangan penilai. Jika beberapa pengalaman penilai diambil kira, rumus min skor adalah:

$$\text{Min skor}_{\text{PK}} = \frac{\sum_{i=1}^n \left(P_i \cdot \prod_{j=1}^m K_{ji} \right)}{\sum_{i=1}^n \left(\prod_{j=1}^m K_{ji} \right)}, \quad (2)$$

dengan $K_{ji} \in [0.1]$, merupakan pengukuran pengalaman penilai i dengan j adalah bilangan jenis pengalaman. Seterusnya jika pengalaman penilai dan pengesanan ralat diambil kira di dalam proses penilaian, rumus bagi min skor yang digunakan adalah:

$$\text{Min skor}_{\text{PKS}} = \frac{\sum_{i=1}^n \left(P_i \cdot \frac{S_i}{k} \cdot \prod_{j=1}^m K_{ji} \right)}{\sum_{i=1}^n \left(\frac{S_i}{k} \cdot \prod_{j=1}^m K_{ji} \right)}, \quad (3)$$

dengan, S_i merupakan bilangan pengesanan ralat yang mendapat respon betul oleh penilai i , dan k adalah jumlah pengesanan ralat di dalam soal selidik. Rumus seterusnya boleh digunakan jika kelayakan penilai diambil kira.

$$\text{Min skor}_{\text{PKSQ}} = \frac{\sum_{i=1}^n \left(P_i \cdot \frac{S_i}{k} \cdot Q_i \cdot \prod_{j=1}^m K_{ji} \right)}{\sum_{i=1}^n \left(\frac{S_i}{k} \cdot Q_i \cdot \prod_{j=1}^m K_{ji} \right)}, \quad (4)$$

dengan Q_i bernilai 0 atau 1. Nilainya 0 jika penilai tidak layak dan 1 jika penilai itu layak.

Dalam masa yang sama analisis faktor penjelajahan (AFJ) dan analisis faktor pengesahan (AFP) di dalam ujian kesahan konstruk bagi instrumen ini telah menghasilkan pekali regresi yang diambil sebagai pemberat kepada metrik. Berdasarkan pemberat metrik, rumus (4) dapat diperkasakan dengan memasukkan pemboleh ubah pemberat metrik ke dalam skor penilai i atau P_i (1) iaitu:

$$P_i = \frac{\sum_{k=1}^m W_k M_k}{\sum_{k=1}^m W_k}, \quad (5)$$

dengan, m ialah jumlah metrik soal selidik, k ialah bilangan metrik, W_k ialah nilai pemberat yang diumpukkan kepada metrik k dan M_k ialah nilai (skala 1 hingga 5) yang diberikan oleh penilai i terhadap metrik k .

Kesimpulannya, rumus kebergunaan dikenal pasti setelah analisis terhadap pemboleh ubah demografi dilakukan dan nilai pemberat metrik ditentukan.

ANALISIS DAPATAN

Penentuan rumus min skor PKM adalah berdasarkan tinjauan hubungan antara pemboleh ubah tak bersandar iaitu metrik demografi dan pemboleh ubah bersandar iaitu min skor kebergunaan. Daripada 650 soal selidik dan PKM yang diedarkan dalam Fasa III, hanya seramai 626 orang (96.3%) memberi maklum balas. Enam kes disingkirkan kerana maklum balas yang tidak boleh dipercayai dikesan. Indeks kebolehppercayaan Cronbach alfa bagi instrumen kebergunaan IPKpkm_4 (64 metrik) tersebut adalah 0.971.

DEMOGRAFI SUBJEK KAJIAN

Responden terdiri daripada 191 orang (30.8%) guru lelaki dan 429 orang (69.2%) perempuan. Guru yang berumur di antara 20 hingga 29 tahun adalah seramai 172 orang (27.7%), berumur 30 hingga 39 tahun seramai 262 orang (42.3%) dan guru yang melebihi 40 tahun seramai 186 orang (30.0%). Daripada segi kelayakan akademik, terdapat 95 orang (15.3%) guru berkelulusan diploma, 503 orang (81.0%) berkelulusan sarjana muda dan 12 orang (1.9%) sahaja mempunyai kelulusan sarjana. Dapatan kajian menunjukkan guru yang mengajar matematik ini seramai 335 orang (54%) mempunyai bidang pengkhususan matematik, seramai 184 orang (29.7%) di dalam bidang Sains dan 90 orang (14.5%) di dalam bidang Sastera.

Guru matematik yang dipilih sebagai responden kebanyakan telah mengajar matematik lebih daripada dua tahun dengan 193 orang (31.1%) telah mengajar antara dua hingga lima tahun, 173 (27.9%) telah mengajar antara enam hingga 10 tahun, dan seramai 198 orang (31.9%) telah mengajar lebih daripada 10 tahun. Seramai 125 orang (20.2%) telah menggunakan aplikasi komputer kurang daripada satu tahun, 312 orang (50.3%) menggunakan komputer antara dua hingga lima tahun, 138 orang (22.3%) telah menggunakan komputer antara enam hingga 10 tahun, dan hanya 44 orang (7.1%) guru telah menggunakan komputer lebih daripada 10 tahun. Antara guru tersebut seramai 504 orang (81.3%) berpengalaman menilai kurang daripada satu tahun dan 112 orang (18.1%) menilai antara dua hingga lima tahun. Hanya 101 orang (16.3%) sahaja yang pernah terlibat dengan pembangunan PKM. Jadual 2 menunjukkan min skor kebergunaan mengikut pemboleh ubah demografi.

Seterusnya hasil ujian hipotesis bagi menentukan pemboleh ubah yang mempengaruhi pemberian skor kebergunaan dibincangkan.

ANALISIS HIPOTESIS KAJIAN

Kesemua hipotesis diuji dengan ujian-t bagi dua pemboleh ubah tak bersandar dan ANOVA satu hala bagi lebih daripada dua pemboleh ubah tak bersandar. Dapatan analisis ditunjukkan dalam Jadual 3 dan Jadual 4. Ujian-t bagi kesamaan min tidak menunjukkan perbezaan yang signifikan di antara guru lelaki dan perempuan, pengalaman membangunkan PKM, pengalaman menilai PKM dengan min skor PKM.

Hasil ujian ANOVA (Jadual 4) tidak dapat membuktikan secara statistik bahawa umur, kelulusan akademik penilai, pengalaman mengajar matematik dan pengalaman menilai PKM tidak mempengaruhi min skor penilaian

JADUAL 2. Frekuensi pemboleh ubah demografi dan min skor kebergunaan

Pemboleh ubah Demografi	Frekuensi	Peratus	Min Skor
Jantina			
Lelaki	191	30.8%	3.7452
Perempuan	429	69.2%	3.7183
Umur			
20-29 tahun	172	27.7%	3.7189
30-39 tahun	262	42.3%	3.7153
40 ke atas	186	30.0%	3.7498
Akademik			
Diploma	95	15.3%	3.7837
Sarjana Muda	503	81.0%	3.7183
Sarjana	12	1.9%	3.6728
Bidang			
Matematik	335	54%	3.7597
Sains	184	29.7%	3.6420
Sastera	90	14.5%	3.7882
Pengalaman mengajar matematik			
0-1 tahun	53	8.5%	3.8377
2-5 tahun	193	31.1%	3.6690
6-10 tahun	173	27.9%	3.7253
> 10 tahun	198	31.9%	3.7545
Pengalaman menggunakan komputer			
0-1 tahun	125	20.2%	3.8009
2-5 tahun	312	50.3%	3.7079
6-10 tahun	138	22.3%	3.7850
> 10 tahun	44	7.1%	3.4632
Pengalaman menilai PKM			
0-1 tahun	504	81.3%	3.7296
2-5 tahun	112	18.1%	3.6949
6-10 tahun	2	.3%	4.5368
> 10 tahun	1	.2%	3.8529
Pengalaman membangunkan PKM			
Ya	101	16.3%	3.6900
Tidak	518	83.5%	3.7597

yang diberikan. Guru yang mengajar matematik yang mempunyai bidang pengkhususan yang berlainan menunjukkan ada perbezaan yang signifikan ($F = 4.286$ dan $p < 0.05$) dalam menilai PKM. Ujian ini juga menunjukkan terdapat perbezaan yang signifikan antara min skor penilaian dengan bidang pengkhususan guru ($F = 4.286$ dan $p < 0.05$) dan pengalaman menggunakan komputer ($F = 6.144$ dan $p < 0.05$).

JADUAL 3. Analisis ujian-t min skor dengan pemboleh ubah demografi

Sumber	t	Sig.
Jantina	0.631	.0529
Membangunkan PKM	-0.742	0.458
Pengalaman menilai PKM	2.017	0.110

JADUAL 4. Analisis Ujian ANOVA min skor dengan pemboleh ubah demografi

Pemboleh ubah	F	Sig.
Umur	0.299	0.742
Kelulusan akademik	0.791	0.454
Pengalaman mengajar matematik	2.011	0.111
Pengalaman menilai PKM	2.017	0.110
Bidang Pengkhususan	4.286	0.014
Pengalaman guna komputer	6.144	0.000

ANALISIS RALAT PENILAIAN

Untuk menjalankan analisis terhadap ralat penilaian, penyelidik telah meletakkan empat metrik sebagai pengesanan ralat di dalam IPK_{pkm_4} iaitu, metrik b27 (Jawapan betul dipaparkan setelah beberapa kali pelajar memberi jawapan salah), metrik e52 (kalkulator disediakan), metrik e53 (ada bahan rujukan tambahan dimasukkan) dan metrik e54 (glosari disediakan). Semua jawapan sepatutnya menunjukkan persetujuan muktamat. Bagi PKM yang dinilai metrik b27, e52 dan e53 sepatutnya dijawab 'tidak setuju' atau 'sangat tidak setuju'. Selain daripada jawapan ini dianggap ralat penilaian telah dilakukan. Bagi metrik e54 jawapan yang diberi sepatutnya adalah 'setuju' atau 'sangat setuju', jawapan lain menunjukkan penilai telah melakukan ralat.

Ujian-t ($t = 7.865$ dan $p < 0.05$) menunjukkan terdapat perbezaan yang signifikan di antara penilai yang melakukan ralat dengan penilai yang tidak melakukan ralat di dalam pemberian skor. Daripada Jadual 2, mereka yang melakukan ralat cenderung untuk memberikan min skor yang tinggi berbanding guru yang tidak melakukan ralat dalam penilaian.

PEMBERAT METRIK

Analisis faktor penjelajahan (AFJ) (Nunnally 1978) dan analisis faktor pengesanan (AFP) (Byrne 2001)

menghasilkan model berstruktur (Sahari et al. 2007). Model padanan telah membuktikan yang faktor pendam Kebolehgunaan (BOLEHGUNA), Kefungsian (FUNGSI) dan Kecekapan (CEKAP) mengukur faktor pendam Kebergunaan (BERGUNA). Terdapat tiga pemboleh ubah pendam secara berperingkat. Peringkat pertama terdapat lapan faktor (Kemudahoperasian (MUDOP), Menyenangkan (SNANG), Kawalan Pengguna (KAWP), Penerangan (TERANG), Pengukuhan (KUKUH), Penilaian (NILAI), Ketepatan (TEPAT) dan Bahan Sokongan Pembelajaran (CEKAP). Peringkat kedua terdapat tiga faktor (Kebolehgunaan, Kefungsian dan Kecekapan). Peringkat ketiga terdapat satu faktor pendam iaitu Kebergunaan Perisian Kursus Matematik. Model ukuran (Rajah 2) menunjukkan model padanan peringkat ketiga. Indeks padanan ($\chi^2/dk = 1.889$; RMR = 0.026; GFI = 0.916; CFI = 0.957; RMSEA = 0.039) menunjukkan bahawa model padanan yang baik telah dihasilkan (Byrne 2001).

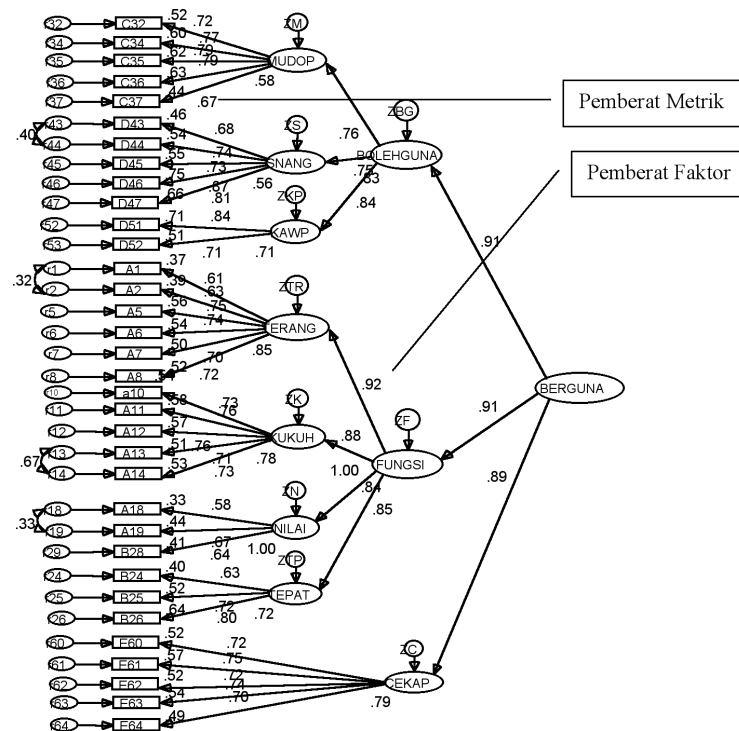
Nilai pemberat regresi piawai (pemberat metrik) bagi setiap metrik kepada faktor masing-masing ditunjukkan melalui Rajah 2. Pemberat faktor Kemudahoperasian 0.67 menunjukkan apabila metrik C37 berubah satu unit, faktor Kemudahoperasian berubah 0.67 unit. Bagi faktor Kemudahoperasian, metrik C36, C35, dan C34, merupakan indikator yang paling utama. Begitu juga dengan faktor Menyenangkan, indikator paling utama adalah metrik D46 dan D47 dan seterusnya.

PERBINCANGAN

Berdasarkan analisis di atas terdapat tiga pemboleh ubah iaitu bidang pengkhususan (Q), pengalaman menggunakan komputer (K) dan ralat penilaian mempengaruhi penilaian PKM. Ini menunjukkan guru matematik yang berpengalaman menggunakan komputer, dengan bidang pengkhususan matematik, dan mereka yang menilai dengan berhati-hati tanpa ralat, memberikan pendapat yang lebih objektif terhadap kualiti kebergunaan PKM yang dinilai. Oleh itu rumus yang sesuai bagi menentukan min skor kebergunaan penilai adalah:

$$\text{Min skor}_{\text{KSQ}} = \frac{\sum_{i=1}^{i=n} \left(P_i \cdot K_i \cdot \frac{S_i}{k} Q_i \right)}{\sum_{i=1}^{i=n} K_i \frac{S_i}{k} Q_i} \quad (6)$$

dengan P_i ialah skor penilai i , K_i ialah pengalaman menggunakan komputer penilai i (dikira berdasarkan pengalaman dan tahun menggunakan suatu produk perisian), S_i ialah bilangan pengesanan ralat yang mendapat respon betul oleh penilai i , k ialah jumlah pengesanan ralat yang digunakan oleh penyelidik, $Q_i = 1$: Jika bidang pengkhususan/opsyen penilai ialah matematik dan $Q_i = 0$: Jika bidang pengkhususan penilai bukan matematik. Penyelidik telah mengumpukkan skor pengalaman penilai berdasarkan graf pengalaman Xenos dan Christodoulakis (1995a) seperti Jadual 5.



RAJAH 2. Model Padanan

JADUAL 5. Skor pengalaman penilai

Pilihan jawapan	Skor penilai
0.0 – 1.0 tahun	0.6
2.0 – 5.0 tahun	0.8
6.0 – 9.0 tahun	0.9
Lebih daripada 10 tahun	1.0

Setelah AFP dilakukan, metrik atau indikator penilaian mempunyai pemberat seperti yang ditunjukkan dalam Rajah 2. Oleh itu rumus min skor kebergunaan dikemas kini dengan menambah pemboleh ubah pemberat metrik ke dalam skor kebergunaan penilai i , P_i (rumus 5). Oleh itu daripada rumus 5 dan 6, min skor kebergunaan PKM adalah:

$$\text{Min skor kebergunaan} = \frac{\sum_{i=1}^{i=n} \left(\frac{\sum_{j=1}^m M_j W_j}{\sum_{j=1}^m W_j} K_i \cdot \frac{S_i}{k} Q_i \right)}{\sum_{i=1}^{i=n} K_i \frac{S_i}{k} Q_i} \quad (7)$$

Dengan, m ialah jumlah metrik soal selidik, k ialah bilangan metrik, W_k ialah nilai pemberat yang diumpukkan kepada metrik k dan M_k ialah nilai (skalar 1 hingga 5) yang diberikan oleh penilai i terhadap metrik k .

Analisis menggunakan ujian-t ($t = 31.783$ dan $p < 0.005$) menunjukkan terdapat perbezaan yang signifikan antara min skor menggunakan rumus lazim (rumus 1) dengan skor min yang dikira menggunakan rumus daripada hasil kajian (rumus 7).

RUMUSAN DAN IMPLIKASI KAJIAN

Kertas ini telah membincangkan pembentukan rumus min skor kebergunaan berdasarkan faktor penilai. Dapatan analisis mencadangkan supaya penyelidik memberi pemberat kepada penilai mengikut bidang pengkhususan guru matematik, pengalaman menggunakan komputer dan ralat penilaian. Daripada dapatan kajian kesahan IPKpkm, setiap metrik telah diumpukkan dengan pemberat tertentu. Oleh itu rumus min skor kebergunaan mengambil kira pemberat dan ketiga-tiga pemboleh ubah tersebut. Dapatan ini menunjukkan guru matematik yang berpengalaman menggunakan komputer, dengan bidang pengkhususan (opsyen) matematik, dan mereka yang menilai dengan berhati-hati tanpa ralat memberikan pendapat yang lebih objektif terhadap kualiti kebergunaan PKM yang dinilai.

Hasil kajian ini boleh digunakan oleh pihak Kementerian Pelajaran atau Institusi Pendidikan sebagai panduan untuk mengukur kesesuaian sebuah PKM sebelum diguna oleh pelajar. Rumus yang dibangunkan dapat mengira skor penilaian oleh sekumpulan penilai khususnya guru matematik dengan mengambil kira latar belakang bidang pengkhususan mereka, pengalaman menggunakan komputer, ralat yang mereka lakukan serta pemberat metrik. Rumus ini tidak mengambil kira sumbangan penilaian yang dilakukan oleh guru matematik yang bukan bidang (opsyen) matematik. Disebabkan kurang pengalaman dan pengetahuan tentang subjek matematik maka kewibawaan guru dalam penilaian dipertikaikan. Ini bersesuaian dengan hasil kajian oleh Marks (1990) yang mendapati tahap pengetahuan pedagogi dan kandungan matematik guru

yang mempunyai pengetahuan matematik yang tinggi dan berpengalaman adalah baik berbanding dengan guru yang kurang pengetahuan matematik (bukan opsyen tetapi mengajar matematik). Seterusnya metrik berpemberat yang dihasilkan boleh digunakan oleh pembangun PKM sebagai panduan untuk menentukan elemen mana yang perlu diberi perhatian. Pembangun juga boleh menggunakan rumus ini dalam penilaian formatif dan sumatif semasa proses pembangunan PKM.

Sebagai kajian lanjutan, rumus ini akan digunakan untuk menghasilkan satu sistem pengukuran kebergunaan sebuah PKM dalam talian. Sistem tersebut boleh digunakan oleh sekumpulan penilai secara serentak dan skor penilaian dapat dijana secara automatik.

RUJUKAN

- Albion, P.R. & Gibson, I.W. 1998. Interactive multimedia and problem based learning: Challenges for instructional design. In *Educational Multimedia and Hypermedia 1998*, edited by T. Ottman & I. Tomek, 117-123. Charlottesville, VA: Association for the Advancement of Computing in Education.
- Atkins, J. 2005. The Association Between the Use of Accelerated Math and Students' Math Achievement. Tesis Doktor Falsafah. East Tennessee State University.
- Byrne, B.M. 2001. *Structural Equation Modeling with AMOS: Basic Concepts Applications, and Programming*. Mahwah, NJ: Lawrence Erlbaum.
- Cook, J. 2002. Review of MathAid Precalculus v15.33. *MSOR Connections* 2(4): 41-44.
- Diaz, P. 2003. Usability of Hypermedia Educational e-Books. *D-Lib Magazine*. 9(3). (atas talian) <http://www.dlib.org/dlib/march03/diaz/03diaz.html>. (04/2/2005). doi: 10.1045/march2003-diaz
- Hawkes, T. 2000. Review of algebra interactive. *MSOR Connections* 0(2): 27-30.
- Jamalludin Harun & Zaidatun Tasir. 2003. *Multimedia dalam Pendidikan*. Kuala Lumpur: PTS Publications & Distributors Sdn. Bhd.
- Kimmins, D. 1995. Technology in School Mathematics: A Course for Prospective Secondary School Mathematics Teachers. *Proceedings of the Eight Annual International Conference on Technology in Collegiate Mathematics*. Houston, Texas. (atas talian) <http://archives.math.utk.edu/ICTCM/EP-8/C89/pdf/paper.pdf> (25/2/2006)
- Lawson, D. 2000. Ride the Rollercoaster with the Mathwise Calculus Cluster. *MSOR Connections* (4): 25-30.
- Marks, R. 1990. Pedagogical Content Knowledge From a Mathematical Case to a Modified Conception. *Journal of Teacher Education* 41(3): 3-12.
- Ministry of Education Malaysia. 1997. *Smart School Flagship application: The Malaysian Smart School-A Conceptual Blueprint*. Kuala Lumpur: Global Printers Sdn. Bhd.
- Nielsen, J. 1994. Heuristic Evaluation. In *Usability Inspection Methods*, edited by Nielsen, J. & Mack, R.J. 25-61. New York: John Wiley & Sons.
- Nunnally, J.C. 1978. *Psychometrik Theory*. Ed. Ke-2. New York: McGraw-Hill.
- Papert, S. 1998. Does Easy Do It? Children, Games and Learning. In *Game Developer Magazine*, Issue June, 1998:87-88.
- Quinn, C.N. 1996. Pragmatic Evaluation: Lessons from Usability. Dlm *Proceedings of ASCILITE 96*. A. Christie, P. James & B. Vaughan (pnyt.). Uni SA, Adelaide.
- Quintana, C., Krajcik, J. & Soloway, E. 2001. Issues and Methods for Evaluating Learner-Centered Scaffolding Advanced Learning Technologies. *Proceedings. IEEE International Conference on 6-8 Aug.* 353-356.
- Sahari @ Ashaari, N., Abdul Ghani, A.A. & Selamat, H., Md Yunus, A.S. 2007. Usefulness Evaluation Model of Mathematics Courseware. *5th International Conference on Quality and Reliability, Chiang Mai, Thailand*: 332-336.
- Squires, D. & Preece, J. 1996. Usability and Learning: Evaluating the Potential of Educational Software. *Computing Education* 27(1): 15-22.
- Xenos, M. & Christodoulakis, D. 1995a. Software Quality: The User's Point of View. In *Software Quality and Productivity*, edited by Matthew Lee, 262-272. International Conference on Software Quality and Productivity, Hong Kong: Chapman & Hall.
- Xenos, M. & Christodoulakis, D. 1995b. Evaluating software quality by the use of user satisfaction measurements. In: *Proc. 4th Software Quality Conference, SET, University of Abertay, Dundee*. 181-188.
- Xenos, M., Stavrinoudis, D. & Christodoulakis, D. 1999. CESM: A Perceived Software Quality Assessment Tool. *Proceedings of the 4th International Conference on Software Process Improvement*. Research into Education and Training, INSPIRE'99. Heraklion, Crete, Greece. 155-164.
- Yerushamly, M. & Schwartz, J.L. 1999. A Procedural Approach to Exploration in Calculus. *International Journal of Mathematical Education in Science and Technology* 30(6): 903-914.
- Yushau, B., Bokhari, M.A. & Wessels, D.C.J. 2004. Computer Aided Learning of Mathematics: Software Evaluation. *Mathematics and Computer Education* 38(2): 165-182.

Noraidah Sahari @ Ashaari, Hairulliza Mohd Judi
Pusat Teknologi Maklumat
Fakulti Teknologi dan Sains Maklumat
Universiti Kebangsaan Malaysia
43600 Bangi, Selangor D.E.
Malaysia

Abdul Azim Abdul Ghani, Mohd Hasan Selamat
Fakulti Sains Komputer dan Teknologi Maklumat
Universiti Putra Malaysia
43400 Serdang, Selangor D.E.
Malaysia

Aida Suraya Md Yunus
Fakulti Pengajian Pendidikan
Universiti Putra Malaysia
43400 Serdang, Selangor D.E.
Malaysia

*Pengarang untuk surat-menyurat; email: nsa@ftsm.ukm.my

Diserahkan: 13 Oktober 2008
Diterima: 29 Januari 2010